

Un exemple d'ambiguïté du geste co-verbal en situation de communication homme-machine et ses conséquences sur les analyses sémantiques et pragmatiques (DRAFT)

Frédéric Landragin

CNRS – Laboratoire Lattice

1, rue Maurice Arnoux

92120 Montrouge

Dans le but de mieux cibler la réalisation de systèmes de dialogue homme-machine, une des étapes consiste à mettre des sujets en situation simulée de communication et à enregistrer leur comportement, à l'aide d'une caméra vidéo et/ou en s'intéressant aux traces captées par les dispositifs d'interaction (microphone, écran tactile, souris). Lorsque le dialogue met en jeu une scène visuelle et autorise les gestes 2D effectués sur écran tactile, on observe essentiellement des gestes co-verbaux déictiques, c'est-à-dire que l'utilisateur pointe les objets et les lieux qu'il mentionne verbalement, suivant en cela le classique paradigme du « *mets ça ici* » de Bolt (1980).

La simulation décrite par Wolff (1999) a conduit à l'enregistrement d'un corpus dans lequel on trouve quelques très rares exemples de gestes co-verbaux non déictiques, en particulier l'énoncé « *mets ça ici* » accompagné d'une trajectoire gestuelle 2D en forme d'arc de cercle. Cette trajectoire peut être considérée comme la matérialisation de la nécessaire transition entre la désignation d'objet et la désignation de lieu. Dans ce cas seules les extrémités de la courbe sont utilisées lors des analyses sémantiques : le point (x_1, y_1) puis l'objet ou le groupe d'objets présent en ce point ou dans un voisinage immédiat sont unifiés avec « *ça* », et le point (x_2, y_2) est unifié avec « *ici* ». Landragin (2004) propose une formalisation des critères de la Gestalt pour intégrer la notion de groupe perceptif à l'algorithme chargé d'identifier à partir du point le ou les objets susceptibles d'être désignés.

La trajectoire gestuelle peut également être considérée comme la combinaison de ces deux désignations avec un geste co-verbal illustratif apportant une caractéristique de l'action de déplacement, à savoir le chemin (ou points de passage) à suivre. Dans ce cas, la trajectoire est analysée d'un point de vue temporel (courbe produite de manière régulière, sans point d'arrêt significatif) et d'un point de vue structurel (arc de cercle, cf. Bellalem & Romary (1996)), avant d'être unifiée à « *déplace* », c'est-à-dire d'être interprétée comme un chemin.

L'intérêt de cette hypothèse repose sur la couverture de l'application et sur les capacités de l'algorithme de résolution des références aux actions : si l'application comporte les deux primitives 'déplacer(objet,lieu)' et 'déplacer(objet,lieu,chemin)', on a une ambiguïté qui n'apparaît pas dans le cas où seule la première primitive est fonctionnelle, c'est-à-dire dans le cas où les points de passage n'interviennent pas dans un déplacement (le déplacement consiste à faire disparaître l'objet de sa position initiale et à le faire réapparaître aussitôt à sa nouvelle position). En fait, on peut considérer qu'il y a ambiguïté même si l'application ne comporte que la première primitive. En effet, le module de résolution des références aux actions peut prendre à sa charge le traitement d'une représentation sémantique incluant le chemin en tant que liste de n points de passage, en n appels de la primitive considérée : 'déplacer(objet,lieu₁)', 'déplacer(objet,lieu₂)', etc., ce qui simule un déplacement avec

points de passage. Cette procédure requiert néanmoins pour la résolution de la référence aux actions des capacités d'interprétation non négligeables.

Si de plus l'application fait la distinction entre l'action de 'déplacer' qui implique la pré-existence de l'objet à déplacer, et l'action de 'mettre' qui implique au contraire un nouvel objet (« *mets une fenêtre ici* »), on peut également considérer que la présence du geste, déictique et/ou illustratif, présuppose que l'objet existe et que le « *mets* » énoncé verbalement doit s'unifier avec l'un des 'déplacer' de l'application. Dans le cas de « *mets ça ici* », il est vrai que la présence de « *ça* » suffit à ce but, mais dans le cas de « *mets la fenêtre ici* » avec le même geste, l'expression « *la fenêtre* » n'est d'aucun apport vu qu'elle ne présuppose même pas la cooccurrence d'un geste co-verbal.

La simple trajectoire que nous avons décrite regroupe ainsi plusieurs éléments qui vont pouvoir être exploités lors des analyses linguistiques. Bien que les systèmes actuels n'entrent pas encore dans ce genre de considérations, il nous semble important d'identifier ce type de paramètres : le geste co-verbal en situation de communication homme-machine, sans atteindre la diversité du geste conversationnel, reste potentiellement polysémique (son étude a donc tout intérêt à exploiter les classifications faites pour le geste conversationnel).

Références bibliographiques

- Bellalem, N., Romary, L. (1996), Structural Analysis of Co-verbal Deictic Gesture in Multimodal Dialogue Systems, *Progress in Gestural Interaction, Proceedings of Gesture Workshop '96*, University of York, UK.
- Bolt, R.A. (1980), Put-That-There: Voice and gesture at the graphics interface, *Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques*, Seattle, Washington.
- Cosnier J. & Vaysse J. (1997), Sémiotique des gestes communicatifs, *Nouveaux Actes Sémiotiques*, 53.
- Landragin F. (2004), *Dialogue homme-machine multimodal – Modélisation cognitive de la référence aux objets*, Hermès-Lavoisier, Paris.
- Wolff F. (1999), Analyse contextuelle des gestes de désignation en dialogue homme-machine, Thèse de l'Université Henri Poincaré, Nancy.